

NOTABREVE

UN ANÁLISIS DE LA EVALUACIÓN ENTRE IGUALES EN REVISTAS CIENTÍFICAS

PEER REVIEW ANALYSIS ON SCIENTIFIC PAPERS

Rodríguez Alcaide, J.J.

Departamento de Producción Animal. Universidad de Córdoba. Campus de Rabanales. 14014 Córdoba. España. pa1roalj@uco.es

PALABRAS CLAVE ADICIONALES

Revistas científicas. Investigación.

ADDITIONAL KEYWORDS

Scientific journals. Research.

RESUMEN

Los evaluadores de los trabajos científicos influyen de modo decisivo en la promoción de la carrera docente e investigadora del profesor universitario. Los evaluadores juegan un gran papel como promotores de la investigación innovadora y ordenan de modo jerárquico a investigadores y publicaciones científicas. Pero los evaluadores se enfrentan a graves cuestiones de las que deben tomar conciencia, entre las que se señala la probabilidad de cometer errores tipo I, dar por bueno algo que es malo, y tipo II, dar por malo algo que es bueno.

SUMMARY

People who evaluate scientific papers have a great influence on professional carrier of scientific research and professor at university level. They promote innovation and a ranking scale among researchers and editors. However those people can make two sort of errors: type I error, by accepting as good something that is bad and type II error, by rejecting as bad something than is good.

INTRODUCCIÓN

Los evaluadores de los trabajos científicos influyen de modo decisivo en el progreso personal del investigador y determinan nuevas oportunidades de investigación, obtención de fondos para investigar y, por tanto, en el nivel de su retribución (Speck, 1993). Los evaluadores ejercen cierto control sobre la calidad del trabajo presentado a la publicación y con sus críticas desean

mejorarlo. También desean con sus críticas promover una investigación innovadora; animan a la diseminación de dicha innovación y sirven para jerarquizar, bajo el principio de la originalidad, a investigadores, publicaciones e instituciones. En verdad, los evaluadores sirven hoy en día para dar una etiqueta de calidad a la investigación que hacen las academias; es decir, son certificadores de calidad. Pero, si esto es así o se pretende que así sea, se debería analizar y evaluar su actuación (Sheperd, 1995).

Las revistas científicas aducen que como editoras buscan una selección adecuada de evaluadores, que además actúen como pares que se desconocen entre sí, y a quienes se le presume ausencia de favoritismo y de ideología, que respondan a la revista con rapidez enviando evaluaciones bien fundadas y llenas de competencia. Obviamente estos principios son una utopía, será difícil no reconocer deficiencias en su satisfacción y cumplimiento (Burham, 1990).

PROBLEMÁTICA Y CRITERIOS PARA LA EVALUACIÓN

¿CUÁLES SON LOS PROBLEMAS A LOS QUE SE ENFRENTAN LOS EVALUADORES?

El papel clave del evaluador

Es decisivo para un investigador y para un departamento o grupo de investigación

Recibido: 8-3-06. Aceptado: 21-3-06.

Arch. Zootec. 57 (219): 333-336. 2008.

publicar sus trabajos. De ello depende su promoción, su salario, la obtención de fondos e incluso en el caso de la persona hasta su salud mental. Pero también depende la reputación de los departamentos y de la Universidad, incluso. Dado que los intereses son tan grandes no se debe soslayar la probable existencia de toda clase de manipulación en relación a los evaluadores. Por eso, para revistas y editores se hace crucial el disponer de evaluadores imparciales y que jueguen limpio (Elliot, 1998).

La impedimenta que deben llevar los evaluadores

Todo evaluador debería atender, al menos, a los siguientes fines:

- Lograr satisfacer niveles predeterminados de calidad (Laband, 1990). Este objetivo tiene tres caras. La primera de ellas es minimizar errores tipo I y tipo II; es decir, que no sean aceptados trabajos defectuosos pero que tampoco sean rechazados trabajos que no tienen defectos. La segunda es prevenir el plagio para proteger los derechos de propiedad y la tercera que, en verdad, publicar en una revista de prestigio sea un sello de excelencia.

- Mejorar la calidad de los manuscritos.
- Promover nuevos paradigmas de modo imparcial.

Dada la nobleza de estos fines, los evaluadores deberían ser imparciales y justos; es decir, que cada investigador debería ser juzgado por el mérito de sus ideas y no por su posición académica, puesto que desempeña, número de publicaciones, etc. (Hamermesh, 1994). Los evaluadores deberían potenciar más que impedir el progreso de la ciencia. Este pronunciamiento no es una cuestión menor, porque la violación de la justicia es una de las mayores ofensas a una conducta científica apropiada.

El dilema del evaluador: minimizar errores tipo I y tipo II

El objetivo de un evaluador es con alto nivel de seguridad identificar la calidad de

cada trabajo; es decir, su validación o el otorgamiento de validez para un fin. La validez no es otra cosa que la completa correlación entre la calidad intrínseca del trabajo y la validación que de esa calidad hace el evaluador (Kraemer, 1991). La validez es la justa correspondencia entre calidad real y calidad enjuiciada y depende del modelo cognitivo ideal que posea el evaluador respecto del trabajo a enjuiciar.

Si los evaluadores fuesen eficientes y su control de calidad también, deberían promocionar aquellas innovaciones que fuesen capaces de superar el filtro diferencial bueno-malo. Revelar la maldad de un trabajo de investigación es una tarea laboriosa y laberíntica, pues la tarea del evaluador se basa sobre la interpretación de las señales, y, por consiguiente, está sometida al riesgo de cometer errores tipo I y tipo II; es decir, aceptar como bueno algo que es malo y rechazar un trabajo que es bueno por considerarlo malo.

No debemos olvidar que existe una transacción entre los dos tipos de errores, pues si un error es grande habrá poca probabilidad de cometer el otro, que será pequeño. Si los editores exigen a sus evaluadores minimizar el error tipo I están al mismo tiempo induciendo a una alta probabilidad de aparición del error tipo II, rechazando trabajos buenos y viceversa (Cole, 1991); en este caso, para no rechazar ningún trabajo bueno, la probabilidad de cometer un error de tipo I se maximiza y apenas se rechazan malos trabajos. Esto es un drama para el editor, pues el contenido de su revista está condicionado por errores tipo I y tipo II. Al editor le preocupa mucho cometer un error de tipo II y por eso en algunas ciencias se rechazan pocos trabajos como en matemáticas, física y química (Hargens and Herting, 1990). Sin embargo, cuando los errores tipo I son más fáciles de identificar que los errores tipo II, los editores se alejan y minimizan el error de dejar pasar un trabajo malo como si fuese bueno. Así pasa en revistas de humanidades y ciencias sociales.

UN ANÁLISIS DE LA EVALUACIÓN ENTRE IGUALES EN REVISTAS CIENTÍFICAS

¿CUÁL ES LA RELACIÓN ENTRE VALIDEZ Y FIABILIDAD?

Los evaluadores deben ser considerados en términos de su fiabilidad, que no es otra cosa que la correlación entre la valoración de calidad que de un trabajo hacen varios evaluadores, debiéndose notar que una elevada fiabilidad puede coexistir con una baja validez (Popper, 1959); es decir, no hay correspondencia entre calidad intrínseca y calidad evaluada y ello sucede cuando los dos evaluadores coinciden pero los dos están equivocados. Igualmente una elevada validez puede existir con una baja fiabilidad; es decir, cuando ambos evaluadores discrepan entre sí. La mayoría de los estudios ponen al descubierto una fiabilidad muy débil entre evaluadores. Si los editores tienen pánico a cometer el error tipo I, ocurre que la fiabilidad entre evaluadores es elevada, pues el editor da más credibilidad al informe negativo. Pero esto impide que se publiquen muchos trabajos que tienen méritos para ser impresos o situados en la red.

¿LA EVALUACIÓN ES UNA SEÑAL DE EXCELENCIA?

Mi sensación es que la publicación de un trabajo no es señal de disseminación de hallazgos, pues las nuevas ideas antes de ser publicadas se han disseminado en seminarios y congresos y mucho más ahora que los profesores tienen páginas digitales en las que colocan sus escritos a disposición de la comunidad académica (Ellison, 2002). En realidad publicar hoy en una revista es pasar a formar parte del archivo de las ciencias y no debe extrañar el escaso número de lectores que reflejan.

DISCUSIÓN

¿SE DA APOYO A LA INVESTIGACIÓN INNOVADORA?

¿Si evita cometer errores de tipo I, está impidiendo la innovación?

Ya se ha dicho. Cuando los evaluadores

y editores tratan de minimizar el error de tipo I, maximizan al mismo tiempo la comisión del error tipo II, que discrimina trabajos innovadores, cuya novedad es reconocida por un número bajo de evaluadores. Si el editor rechaza un trabajo cuando recibe una evaluación negativa de un solo evaluador, la innovación tiene enormes dificultades de ser impresa o digitalizada en una revista. Cuando el investigador conoce estas condiciones se concentra en cumplir las normas que aparecen en la contraportada, exigidas para enviar artículos y no presta atención a lo que la innovación representa. Los investigadores se dedican a pulir su trabajo en lugar de desarrollar sus nuevas visiones (Crandall, 1991). Los editores y evaluadores lo niegan pero deben saber que si minimizan el error tipo I están maximizando el error tipo II. Minimizar el error tipo I conduce a veces a favoritismo y a sesgos enormes en la evaluación.

¿Si evita cometer errores tipo I es proclive al favoritismo?

El prestigio del autor y las relaciones personales son señales de calidad para evaluadores anónimos. Muchas revistas, si el autor es desconocido ni siquiera pasan el trabajo al evaluador (Hodgson and Rothman, 1999). No es lo mismo para muchas revistas que el trabajo proceda de un autor cuando pertenece a una universidad de segundo orden, Coupé (2000). Se cuenta de un profesor de Dakota del Norte que envió 20 artículos que nunca fueron publicados en revistas principales, pero cuando se trasladó a Harvard, aquellos mismos artículos encontraron su lugar en las mismas revistas que los habían rechazado (Rosenthol, 1982).

El evaluador anónimo también utiliza las mismas señales del editor en relación a la identidad del grupo y autores, que se obtienen de las referencias bibliográficas y líneas de investigación. Esto refuerza el sesgo del editor hacia el favoritismo, que se refuerza por el autor que usa el inglés como lengua vernácula. El dominio del inglés es

en general cauce de favoritismo, que se puede, no siempre, resolver con una buena traducción.

El sesgo de la confirmación de hipótesis

Existen estudios empíricos que demuestran que se publican más estudios que confirman las hipótesis que los que las niegan. Lo mismo ocurre cuando los resultados se pueden aseverar con altos niveles de seguridad estadística frente a quienes no alcanzan tan elevados niveles de seguridad. Lo mismo ocurre en relación al muestreo grande versus pequeño aleatorio. Esta tendencia es alarmante porque la ciencia avanza más, bajo el principio de falsabilidad de

Popper (1959) que bajo el principio de verificabilidad; una buena revisión aparece en Seide *et al.*, 2005. Debe saberse que los resultados positivos o negativos de una experiencia no son equivalentes cuando se analizan y discuten sus implicaciones lógicas. A pesar de todo ello las revistas siguen publicando más trabajos con resultados que confirman las hipótesis que los trabajos que la niegan.

COROLARIO

¿Sigue usted confiando en que se publiquen sus artículos en revistas de impacto y en el impacto de sus artículos científicos?.

BIBLIOGRAFÍA

- Burham, J.C. 1990. The evolution of editorial peer review. *J. Amer. J. Ass.*, 263: 1323-1329.
- Cole, S. 1991. Consensus and reliability of peer-review evaluations. *Behav. Brain Sci.*, 14: 140-141.
- Coupé, T. 2000. Data on the main regions of journal contributors homepages. ulb.ac.be/~tcoupe/update/journreg.html. Acceso: marzo 2006.
- Crandall, R. 1991. What should be done to improve reviewing. *Behav. Brain Sci.*, 14: 143.
- Elliot, C. 1998. Who's publishing who? The national composition of contributors to some core US and European Journals. *Europ. Econ. Rev.*, 42: 201-206.
- Ellison, G. 2002. Evolving standards for academia publishing: a q-r theory. *J. Polit. Econ.*, 110: 994-1034.
- Hamermesh, D.S. 1994. Facts and myths about refereeing. *J. Econ. Perspect.*, 8:153-163.
- Hargens, L.L. and J.R. Herting. 1990. A new approach to references assessments of manuscripts. *Soc. Sci. Res.*, 19: 1-16.
- Hodgson, G.M. and H. Rothman. 1999. The editors and authors of economic journal: a case of institutional oligopoly?. *The Econ. J.*, 109: F165-186.
- Kraemer, H.C. 1991. Do we really want more reliable reviewers?. *Behav. Brain Sci.*, 14:152-154.
- Laband, D.N. 1990. Is there value-added from the review process in economics? Preliminary evidence from authors. *The Quart. J. Econ.*, 105: 341-352.
- Popper, K.R. 1959. The logic of scientific discovery. Basic Books. N.Y.
- Rosenthal, R. 1982. Reliability and bias in peer-review practices. *Behav. Brain Sci.*, 5: 235-236.
- Seide, C., V. Schmidt and P. Grösche. 2005. The performance of peer review and a beauty contest of referee processes of economics journal. *Estud. Econ. Apl.*, 23: 505-554.
- Sheperd, G.B. 1995. Rejected, leading economist ponder to publications process. Thomas Morton and daughters. Sun Lakes. Arizona.
- Speck, B.W. 1993. Publication peer review. An annotated bibliography. Greenwood Press. Westpoint. Conn and London.